
Lambert et la loi de survie

Lambert and the life table

Marc Barbut, Jean-Marc Rohrbasser and Jacques Veron

**Electronic version**

URL: <http://journals.openedition.org/msh/2970>

DOI: 10.4000/msh.2970

ISSN: 1950-6821

Publisher

Centre d'analyse et de mathématique sociales de l'EHESS

Printed version

Date of publication: 1 September 2005

ISSN: 0987-6936

Electronic reference

Marc Barbut, Jean-Marc Rohrbasser and Jacques Veron, « Lambert et la loi de survie », *Mathématiques et sciences humaines* [Online], 171 | Automne 2005, Online since 10 April 2006, connection on 20 April 2019. URL : <http://journals.openedition.org/msh/2970> ; DOI : 10.4000/msh.2970

DOSSIER « DEMOGRAPHIE MATHÉMATIQUE »

– 6 –

LAMBERT ET LA LOI DE SURVIE¹

Marc BARBUT², Jean-Marc ROHRBASSER³ et Jacques VÉRON⁴

RÉSUMÉ – *J.H. Lambert (1728-1777) s'est illustré par ses nombreux travaux tant en mathématiques pures qu'en applications des mathématiques. Parmi ses domaines d'application figure la démographie.*

Cet article présente un panorama des « modèles » formulés par Lambert pour décrire, en particulier à partir des bulletins de mortalité de la ville de Londres, le processus de mortalité dans des populations stationnaires.

Une hypothèse est proposée quant à la façon dont le principal de ces modèles aurait pu être élaboré.

MOTS-CLÉS – Démographie, Equations de survie, J.H. Lambert, Tables de mortalité, Vie moyenne, Vie probable (syn. Vie médiane).

SUMMARY – Lambert and the Life Table

J.H. Lambert (1728-1777) is renowned for his works in pure and applied mathematics. But very little is known about the “mathematical models” he worked out for demographic data.

Here we present a panorama of his mathematical theory of the so-named “population pyramid” in a stationary population.

Moreover, we suggest one explanation for the genesis of his main population model.

KEYWORDS – Bills of mortality, J.H. Lambert, Mean life, Median life, Population studies, Survival equations.

Entre 1765 et 1772, Johann Heinrich Lambert rédige un ensemble de traités de mathématiques, qu'il nomme *Contributions à l'utilisation des mathématiques et à leur application*⁵, dans lequel il applique notamment celles-ci aux questions de population,

¹ Les auteurs remercient Roland Pressat dont une relecture attentive de leur texte a permis de notamment l'améliorer et de corriger quelques erreurs. Celles qui resteraient sont de leur seul fait.

² CAMS, EHESS, 54 boulevard Raspail, 75006 Paris, mbarbut@ehess.fr

³ INED, 133 boulevard Davout, 75980 Paris Cedex 20, rohrbass@ined.fr

⁴ INED, 133 boulevard Davout, 75980 Paris Cedex 20, veron@ined.fr

⁵ *Beyträge zum Gebrauche der Mathematik und deren Anwendung* durch J. H. Lambert, Berlin, 1765-1772. Les pages sur les questions de population feront l'objet d'une édition critique et d'une traduction française par Jean-Marc Rohrbasser et Jacques Véron, avec une contribution de Marc Barbut, à paraître en 2006 aux éditions de l'INED, collection «*Classiques de l'économie et de la population*».

Les *Beyträge*... sont divisées en trois parties («*Heile*») réunies en trois volumes. La deuxième partie est

plus particulièrement à l'analyse de la mortalité et de la nuptialité ainsi qu'au risque de contracter la petite vérole.

Lorsque Lambert développe ses conceptions sur la population, il dispose, comme source principale, de l'ouvrage récemment paru de Süßmilch, *Die göttliche Ordnung in den Veränderungen des menschlichen Geschlechts*, dont la première version date de 1741, la deuxième, très augmentée, de 1761-62, et la troisième, identique à la précédente, de 1765. Dans ses *Contributions...* Lambert se réfère principalement aux données et aux analyses du pasteur prussien.

Par rapport aux travaux antérieurs, ceux de Lambert manifestent un souci de théorisation et de généralisation. Il a largement recours à l'analyse mais la méthode géométrique demeure pour lui fondamentale.

Dans cet article, nous allons nous concentrer sur l'analyse que Lambert a faite du processus de mortalité.

DONNEES DE MORTALITE ET «REALITE»

Lambert se fonde sur les bulletins de mortalité de Londres de 1728 à 1758, tels qu'il les trouve dans le *Göttliche Ordnung*... de 1761 de Süßmilch et dans *A Collection of the Yearly Bills of Mortality, From 1657 to 1758 inclusive*⁶ Le mathématicien se montre très scrupuleux dans l'usage des données et dans leur retranscription. Dans le même temps, il considère que la régularité des données est une preuve de leur «qualité», c'est à dire de leur adéquation au réel. La nature ne fait, pour Lambert, pas de «sauts» (*Sprünge*) :

[...] dans la dernière colonne où [Süßmilch] indique sur combien d'individus un seul meurt à chaque âge, on voit qu'il se produit des sauts qui s'écartent fort de la loi de la nature⁷

Dans ce même paragraphe, Lambert évalue en quelque sorte la qualité des chiffres de Süßmilch suivant le critère de la régularité :

Ici, les nombres 90, 89, 88, 87, 86 se suivent dans leur ordre naturel. Mais pourquoi, à la 34^e année, cette chute brusque de 86 à 71, et, plus bas, de 70 à 60 ? Ce sont là des inexactitudes [Unrichtigkeiten] qui, dans de telles tables, doivent aussi bien être évitées en elles-mêmes, que, dans le cas particulier où, si l'on ne veut pas avoir recours à des fractions décimales, le calcul doit être fait, non pour 1000, mais pour 10 000 ou 100 000⁸

Les courbes de mortalité, mais également celles de nuptialité, doivent donc, selon lui, être régulières, ce qui l'autorise, le cas échéant, à «lisser» les dites courbes. Lambert est bien conscient des contraintes et des sources de biais liées aux données dont il dispose :

elle-même divisée en deux sections («Abschnitte»). Les considérations sur la population figurent dans le premier volume de 1765 et dans le troisième, de 1772. Les autres développements de Lambert dans ces trois volumes portent notamment sur la mécanique, l'astronomie, la géométrie pratique, l'algèbre, la trigonométrie et la gnomonique.

⁶ *A Collection of the Yearly Bills of Mortality, from 1657 to 1758 inclusive. Together with several other Bills of an earlier Date*, London, Printed for A. Millar in the Strand, 1759.

⁷ *Beyträge* III 1772, IX «Anmerkungen über die Sterblichkeit, Todtenlisten, Geburten und Ehen», [«Observations sur la mortalité, les listes de décès, les naissances et les mariages»], § 19.

⁸ *Ibid.*

Mais comme, à cause du trop petit nombre des observations, les nombres de cette table sont encore trop irréguliers [zu viel irregulär], je vais en ajouter une autre [table] par laquelle on pourra rectifier [berichtigtet] la colonne v.⁹

Utilisant d'abord les bulletins de la mortalité de Londres sur six années (1753-1758) puis sur trente années (1728-1757) et constatant que « Les nombres qui ne sont déduits que de 6 années ne sont pas différents de ceux que l'on tire de 30 années au point que l'on remarque une différence dans la figure », Lambert considère que celle-ci permet de « Se faire d'un coup d'œil une idée générale de la loi de la mortalité à Londres »¹⁰. Il recherche ensuite « Une équation *a posteriori* » telle « Qu'elle ne s'écarte pas notablement du vrai »¹¹.

D'abord en 1765 puis en 1772, Lambert propose des formalisations du processus de mortalité. Les deux équations présentées dans chacun des volumes sont radicalement différentes.

LA PREMIERE EQUATION DE LA SURVIE

Lambert établit son équation de 1765 à partir des données de mortalité concernant la ville de Londres, pour les 6 années de 1753 à 1758 incluses. Ces données sont fournies par Lambert au paragraphe 67 de son ouvrage. Le tableau ci-dessous en est la forme normalisée pour une base de 100.000 naissances vivantes, telle que fournie par Lambert.

Tableau 1. Nombres, réduits par Lambert, des survivants à Londres de 1753 à 1758

Age en années	0	2	5	10	20	30	
Nombre de vivants	100.000	64.136	54.846	51.659	48.831	41.246	

Age en années	40	50	60	70	80	90	100
Nombre de vivants	31.992	22.474	14.146	7.435	2.593	347	13

La forme générale de cette équation est :

$$(1) \quad y(x) = a + bx + cx^2 + dx^3 + ex^4 + fx^5$$

où la lettre x représente le nombre d'années vécues au-delà de l'âge de 45 ans, et $y(x)$ le nombre correspondant de survivants (pour l'étude algébrique de ce polynôme, cf. Annexe II).

Le polynôme (1) est une approximation de la courbe continue, tracée « à main levée »¹² par Lambert en interpolant celles des données qui ne concernent que les âges

⁹ § 131.

¹⁰ § 8.

¹¹ § 9.

¹² *Beyträge* I, 1765, « Theorie der Zuverlässigkeit der Beobachtungen und Versuche » [« Théorie de la fiabilité des observations et des épreuves »], § 67 et sq.

de 40 à 100 ans, avec un pas de 10 ans (cf. Figure 1, «Figure VII» de Lambert) le point A y correspond à l'âge de 45 ans et le point D à 90 ans).

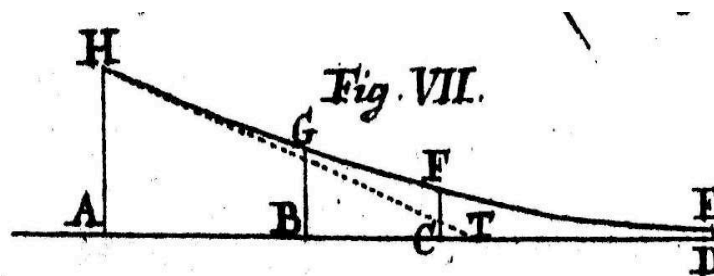


Figure 1. Morceau de la courbe de survie donnée par Lambert en 1765

D'après sa figure, Lambert évalue la population de survivants à l'âge de 45 ans (soit $y(0)$) à 26 950 âmes. Cet âge de 45 ans n'a sûrement pas été choisi au hasard l'examen des données montre en effet que c'est à peu près à cet âge que doit se situer la 2^e inflexion de la courbe, le sens de la courbure ne change plus au-delà.

Le coefficient a de (1) étant ainsi déterminé, il reste à trouver les cinq autres. Lambert ne connaît pas les méthodes de type «moindres écarts» élaborées ultérieurement par les Laplace, Legendre, Gauss ou Cauchy. La première idée qui peut venir à l'esprit, c'est de faire passer la courbe représentative du polynôme (1) par cinq points correspondant aux observations par exemple, pour 50, 60, 70, 80 et 90 ans (i.e. $x = 5, 15, 25, 35$ et 45). On obtient ainsi un système de cinq équations linéaires en b, c, d, e et f qu'il ne reste qu'à résoudre. Mais attention aux erreurs d'arrondi, qui vont s'accumuler. En outre, rien n'assure *a priori* que le polynôme obtenu aura une courbure de sens constant.

Très astucieusement, Lambert choisit une autre voie il va faire en sorte que le polynôme (1) soit tangent à la courbe tracée «à main levée». Cette évaluation, faite graphiquement, le conduit aux pentes de -985,7 pour la tangente en H (i.e., $b = -985,7$ dans (1)), et de -64 en E . Admirable précision.

En faisant passer son polynôme par trois points supplémentaires parmi les observations (en l'occurrence, par ceux correspondant à 60, 70 et 90 ans, $x = 15, x = 25$ et $x = 45$), il est ainsi conduit au système de 4 équations linéaires

$$14\,146 = 26\,950 - 985,7 \cdot 15 + 15^2 \cdot c + 15^3 \cdot d + 15^4 \cdot e + 15^5 \cdot f$$

$$7\,435 = 26\,950 - 985,7 \cdot 25 + 25^2 \cdot c + 25^3 \cdot d + 25^4 \cdot e + 25^5 \cdot f$$

$$347 = 26\,950 - 985,7 \cdot 45 + 45^2 \cdot c + 45^3 \cdot d + 45^4 \cdot e + 45^5 \cdot f$$

$$64 = -985,7 + 2 \cdot 45 \cdot c + 3 \cdot 45^2 \cdot d + 4 \cdot 45^3 \cdot e + 5 \cdot 45^4 \cdot f$$

pour déterminer c, d, e et f (pour le calcul algébrique de la dérivée, se reporter à l'Annexe II).

Seulement, il a commis une erreur, que nous appellerions «détourderie» après avoir correctement écrit qu'en E la courbe «s'approche chaque année de 64 de l'axe des abscisses», il donne, dans toutes les équations qui suivent, la valeur +64 (et non -64) à la pente. Et l'on vérifie immédiatement que les valeurs qu'il fournit pour les coefficients correspondent bien, approximativement, à +64.

Si maintenant l'on résout le même système de quatre équations linéaires, mais en substituant -64 à +64 dans la dernière, on obtient le tableau suivant pour les valeurs des coefficients□

Tableau 2. Valeur des coefficients de la première équation de survie

	Valeurs recalculées	Valeurs données par Lambert
c	+ 11,423445	+ 9,709150
d	- 0,2564623	- 0,0342700
e	+ 0,006007335	- 0,0027017
f	- 0,000035998	+ 0,000066635

Comparons maintenant les observations et les valeurs estimées selon qu'on utilise pour le polynôme (1) les coefficients donnés par Lambert ou les coefficients recalculés (cf. Tableau 3).

Tableau 3. Comparaisons entre données observées et données calculées par Lambert

Age	x	\hat{y} Lambert	$\hat{y} - y$	y observé	$\hat{y} - y$	\hat{y} recalculé
40	- 5	32 123	+ 131	31 992	+ 208	32 200
50	5	22 258	- 216	22 474	- 195	22 279
60	15	14 147	+ 1	14 146	0	14 146
70	25	7 435	0	7 435	0	7 435
80	35	2 320	- 273	2 593	- 20	2 573
90	45	349	+ 2	347	0	347

Que pour $x = 15, 25$ et 45 l'écart soit nul ou quasi nul dans l'un et l'autre cas est normal□ puisque ce sont les valeurs de x qui ont servi au calcul des coefficients. L'un et l'autre de ces ajustements restent très acceptables pour $x = 5$ et $x = - 5$, valeurs qui n'ont pas servi pour la détermination des coefficients. En effet, l'écart aux observations y est inférieur à 1 %. En revanche, l'ajustement par le polynôme (1) n'a plus aucun sens pour les âges inférieurs à 40 ans, de même que pour les âges supérieurs à 90 ans□ le polynôme (1) passe en effet par un minimum pour un x_0 situé entre 44 et 45, et croît ensuite très rapidement (cf. Annexe II). Quant aux valeurs trouvées pour $x = 35$, seule la formule recalculée est acceptable, avec un écart à l'observation inférieur à 1 %. La formule de Lambert donne par contre un écart très important, supérieur à 10 %. Ce seul fait aurait dû alerter le mathématicien et lui indiquer que quelque chose «clochait□.

En 1772, Lambert propose une tout autre équation de la survie, sans établir de rapprochement avec la précédente.

LA DEUXIEME EQUATION DE LA SURVIE

C'est dans la troisième partie des *Contributions...* que Lambert présente une deuxième équation de la survie :

$$y(x) = 10000 \left(\frac{96-x}{96} \right)^2 - 6176 \left(e^{-\frac{x}{31,682}} - e^{-\frac{x}{2,43114}} \right)$$

Ici, x représente l'âge en années (à partir de la naissance).

Cette équation associe trois fonctions (cf. Figure 2), qui conduisent Lambert à comparer le processus de mortalité à des phénomènes d'ordre physique□

L'équation, très simple, est composée d'une parabole et de deux droites logistiques. Le premier membre, parabolique, indiquerait que l'espèce humaine s'éteint à la manière d'un récipient cylindrique qui se vide de son eau. Les deux autres membres ont quelque chose de très analogue au réchauffement et au refroidissement des corps, puisqu'ils sont fondés sur la droite logistique [...]¹³.

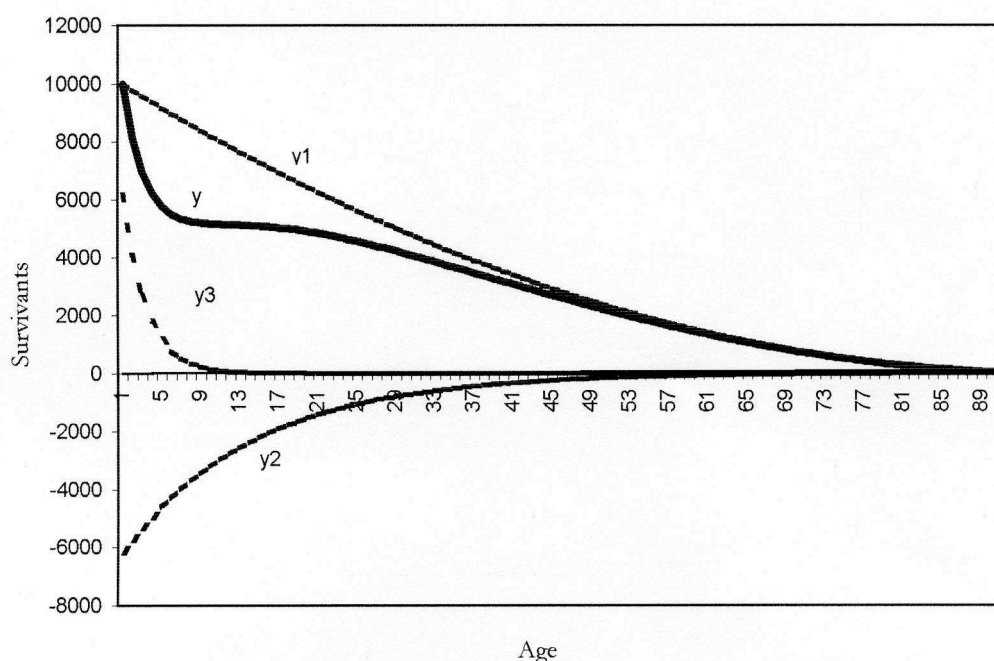


Figure 2. Courbe de mortalité tracée à partir de l'équation de Lambert
($y = y_1 + y_2 + y_3$)

L'application que nous pouvons faire de cette formule aux données de Londres utilisées par Lambert met, au demeurant, en évidence une inversion de deux chiffres dans la première exponentielle□il faut lire 13,682 au lieu de 31,682.

La véritable formule est donc (pour l'étude de la fonction y , se reporter à l'Annexe III) :

¹³ *Beyträge* III, § 10.

$$y(x) = 10000 \left(\frac{96 - x}{96} \right)^2 - 6176 \left(e^{-\frac{x}{13,682}} - e^{-\frac{x}{2,43114}} \right)$$

Le mathématicien avait lui-même reconnu cette inversion dans une lettre adressée aux pères Roberto Gaeta et Gregorio Fontana, les auteurs de la traduction italienne de la *Doctrine of Chances* d'Abraham de Moivre¹⁴. Dans cette lettre, Lambert montrait également son souci de formaliser le processus de mortalité de la manière la plus simple et la plus générale possible□

Berlin, 16 mars 1776

Sur ce que vous me marquez au sujet de la formule qui exprime assez bien la loi de la mortalité de Londres, je l'ai examinée de nouveau, & je vois qu'une faute d'écriture ou d'impression l'a défigurée. L'exposant du second terme au lieu de $-x : 31,682$, doit être $-x : 13,682$. Je vous suis bien obligé de m'y avoir rendu attentif.

Du reste je n'ai pas employé quelque méthode particulière pour trouver cette Formule. Aussi je ne la donne que parce qu'elle répond assez bien aux observations jusqu'à l'âge de 96. ans. J'avertis [e]ncore qu'elle ne doit pas être étendue au-delà de ce terme.

J'aurois plutôt souhaité de trouver une formule simple qui exprimât assez exactement les nombres du §. 23. Mais toutes les méthodes employées jusqu'à présent ont été infructueuses. Elles donnent des formules trop longues ou trop peu exactes. Du reste ces méthodes consistent en ce que je regarde les nombres du §.23. comme des abscisses & des ordonnées d'une ligne courbe. En la construisant sur une grande feuille on voit d'abord qu'elle a deux points d'inflexion contraires. Cela fait que l'équation $\frac{d^2y}{dx^2} = 0$ doit avoir deux racines réelles. Elle n'a point de maximum, ni de

minimum ; cela fait que l'équation $\frac{dy}{dx} = 0$ n'a point de racine réelle, à

moins que ce ne soit la valeur de $x = \text{infini}$. Car la courbe paroît être asymptotique. Du moins il ne sera guère possible d'assigner un âge, qui doive nécessairement terminer la vie. Mais comme ces considérations ne suffisent pas pour déterminer entièrement la nature de cette courbe, quoiqu'elles excluent toutes les lignes courbes qui n'y répondent pas, il s'agit de voir comment cette courbe pourroit du moins à très peu près être représentée par la somme ou la différence des quelques lignes courbes plus connues. La condition qu'elle est asymptotique, demande des lignes logarithmiques, ou des hyperboles du genre

$\frac{a + bx + cx^2 + ec.}{\dot{a} + \dot{a}x + \dot{a}x^2 + ec.}$ qui soient asymptotiques. Avec tout cela, il m'a toujours

paru que cette courbe après l'âge de 70. ans s'approche de la ligne des abscisses comme feroit une parabole qui auroit son sommet au point qui répond à l'âge d'environ 96. ans¹⁵.

¹⁴ L'actuaire Albert Quiquet (*Aperçu historique sur les formules d'interpolation des tables de survie et de mortalité*, Deuxième édition, Paris, Warnier et C^{ie}, 1893, p. 10) mentionne cette inversion et la réponse de Lambert aux pères Gaeta et Fontana□cf. infra).

¹⁵ Roberto Gaeta et Gregorio Fontana, «Discorso preliminare□, *La Dottrina degli Azzardi...* [traduction

Comment Lambert a-t-il donc établi sa seconde équation de la mortalité ? Quelques hypothèses peuvent être avancées. Avec $B = 6\,176$, $\lambda = \frac{1}{13,682}$ et $\mu = \frac{1}{2,43114}$, l'équation de Lambert s'écrit :

$$(2) \quad y(x) = 10\,000 \left(\frac{96-x}{96} \right)^2 - B \left(e^{-\lambda x} - e^{-\mu x} \right)$$

Dans (2), x est toujours l'âge en années, $y(x)$ le nombre de survivants à l'âge x , rapporté à 10 000 naissances vivantes (les mort-nés ont été soustraits).

Les données utilisées pour établir l'expression (2) auraient été, selon Gaeta et Fontana, celles fournies par le recueil *A Collection of the Yearly Bills of Mortality, From 1657 to 1758 inclusive* donnant les nombres de décès par âge, à Londres pour les 30 années de 1728 à 1757¹⁶. Celles-ci ont été résumées par Süßmilch dans sa Table X (cf. Annexe I) que Lambert dit avoir utilisée.

Quant à la méthode suivie par Lambert pour établir sa formule (2), voici ce qu'en disent les pères R. Gaeta et G. Fontana :

[...] après des tentatives et des essais nombreux et variés faits à tâtons (cela a dû vraisemblablement se passer, mais l'auteur a entièrement supprimé de son texte la voie qu'il a empruntée) on trouve finalement la belle équation [...] ¹⁷.

Cette « belle équation », c'est l'équation (2). Sur cette méthode par approximations utilisée par Lambert, une *première hypothèse* est que le terme quadratique initial $10\,000 \left(\frac{96-x}{96} \right)^2$ a été choisi *a priori*.¹⁸

D'autre part, dans ce que nous appellerions aujourd'hui la modélisation de la mortalité, ce premier terme correspond au fait que « l'espèce humaine vient à décéder [...] graduellement, exactement comme on voit en hydraulique un vase cylindrique plein d'eau se vider progressivement [...] ». ¹⁹ Or, cette loi de l'hydraulique était évidemment connue de Lambert, et l'analogie entre l'écoulement de la vie humaine et celui d'un fluide ne manque pas de pertinence.

Si le terme complémentaire est désigné par $u(x)$, on a

$$(3) \quad y(x) = 10\,000 \left(\frac{96-x}{96} \right)^2 - u(x)$$

italienne de *The Doctrine of Chances* d'Abraham de Moivre], Milan, 1776, p. lvi-lvii. Ce texte est traduit et reproduit par Pierre Crépel dans le volume *Arithmétique politique française*, Paris, Ined, 2003, p. 299.

¹⁶ Ces données sont résumées par la table X de Süßmilch (cf. Annexe I) et c'est vraisemblablement cette seule table X que Lambert a utilisée.

¹⁷ Gaeta et Fontana, *op. cit.*, p. 297.

¹⁸ On peut noter que, dans la table du paragraphe 461 de l'*Ordre divin* de Süßmilch, il reste, sur 1000 naissances, deux survivants à 96 ans et un seul survivant à 97 ans.

¹⁹ Gaeta et Fontana, *op. cit.*, p. 298.

$y(x)$ résulte des observations — on calcule aisément ses valeurs au moyen de la table X de Süßmilch dernière colonne (Annexe I *infra*). Donc $u(x)$ observé se calcule par (3') —

$$(3') \quad u(x) = 10\,000 \left(\frac{96-x}{96} \right)^2 - y(x)$$

On obtient ainsi le Tableau 4 et la Figure 3. (Il est possible que, le calcul ayant été fait « à la main », des erreurs de quelques unités, donc négligeables, se soient introduites dans le Tableau).

Tableau 4. Calcul de $u(x)$ en fonction de l'âge x .

x	2	5	10	20	30	40	50	60	70	80	90
$u(x)$	3218	3485	2865	1417,5	646,5	283	146	56,3	13,5	7,8	-0,1

Représentons graphiquement ces données :

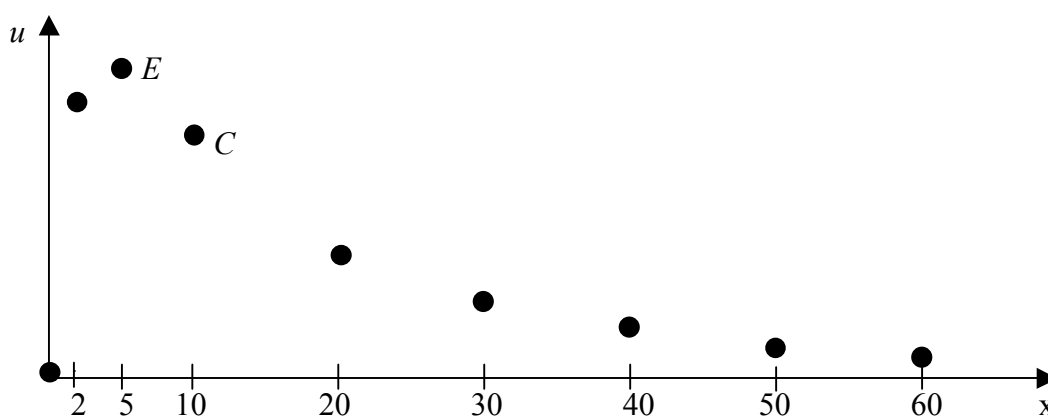


Figure 3. Valeurs de $u(x)$ en fonction de l'âge x .

Une courbe théorique ajustée à ces données est de la forme présentée dans la Figure 4.

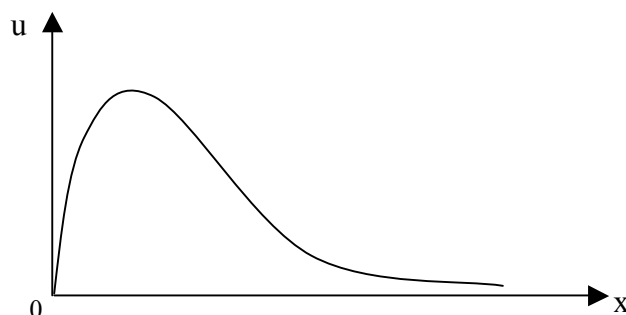


Figure 4. Allure de la courbe ajustant les valeurs de $u(x)$ en fonction de x .

Or, Lambert connaissait bien une famille de courbes ayant cette forme. Ce sont celles d'équation —

$$(4) \quad f(x) = B \left(e^{-\lambda x} - e^{-\mu x} \right)$$

qui interviennent en théorie de la chaleur.

Lisons en effet de nouveau les pères Gaeta et Fontana□

... [Ce terme complémentaire de (2)] *montre que la mortalité suit à cet égard, à peu de chose près, l'échelle et la graduation qu'il y a justement dans le réchauffement et le refroidissement des corps, graduation qui a justement pour base la courbe Logistique, comme l'a démontré le même M. Lambert dans le second tome des Actes Helvétiques*²⁰.

Ajoutons que «*De même M. Lambert*□ est très fier de savoir estimer les paramètres d'une fonction au moyen d'une courbe « *tracée à main levée*□. La méthode qu'il expose pour établir sa première formule (cf. *supra*) repose notamment sur l'estimation de la pente aux points d'inflexion. Dès lors, la suite devient claire.

L'étude des fonctions de la forme (4) montre (*ln* désignant le logarithme népérien) que le maximum est atteint pour $x_0 = \frac{1}{\mu - \lambda} \ln\left(\frac{\mu}{\lambda}\right)$. Il y a un point d'inflexion *c* pour l'abscisse $2x_0$. La tangente à l'origine a pour pente : $f'(0) = B(\mu - \lambda)$. La courbe se présente alors ainsi (cf. Figure 5).

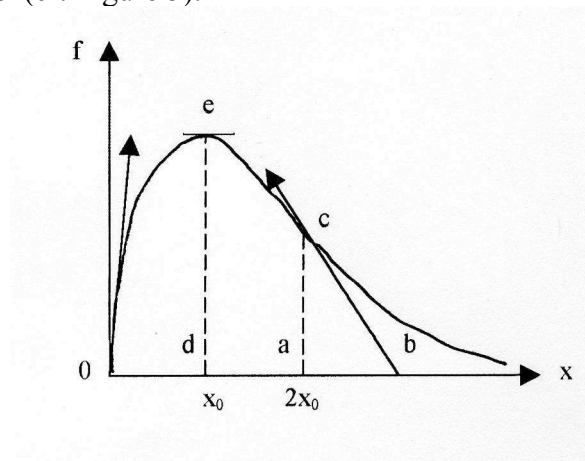


Figure 5. Représentation graphique du deuxième terme.

Les expressions de $f(x_0)$, $f(2x_0)$ et $f'(2x_0)$ font intervenir des fonctions transcendantes de λ et μ . Mais on démontre aisément les deux relations suivantes (cf. Annexe IV)□

$$(5) \quad \lambda + \mu = \frac{f(2x_0) f'(0)}{f^2(x_0)}$$

$$(6) \quad \frac{1}{\lambda} + \frac{1}{\mu} = \frac{\lambda + \mu}{\lambda\mu} = \frac{f(2x_0)}{|f'(2x_0)|}$$

²⁰ Gaeta et Fontana, *op. cit.*, p. 298.

On remarquera d'ailleurs que le rapport $f(2x_0)/|f'(2x_0)|$ n'est autre que la longueur ab de la figure 5.

Posons $\frac{1}{\lambda} = \lambda', \quad \frac{1}{\mu} = \mu'.$

Il en résulte $\lambda + \mu = \frac{1}{\lambda'} + \frac{1}{\mu'} = \frac{\lambda' + \mu'}{\lambda' \times \mu'}$ et $\lambda \times \mu = \frac{\lambda' + \mu'}{\lambda + \mu}$

D'où

$$(6') \quad \lambda' + \mu' = \frac{f(2x_0)}{|f'(2x_0)|} = ab = \alpha$$

$$(5') \quad \lambda \times \mu = \frac{\lambda' + \mu'}{\lambda + \mu} = (\lambda' + \mu') \cdot \frac{f^2(x_0)}{f(2x_0)f'(0)} = \frac{ab \times (de)^2}{ac \times f'(0)} = \beta$$

λ' et μ' sont alors les racines de l'équation du 2^e degré

$$t^2 - \alpha t + \beta = 0$$

On obtient ainsi λ et μ , puis B par

$$B = \frac{f'(0)}{\mu - \lambda}$$

Si l'on revient aux données empiriques du Tableau 3 et de la Figure 3, on observe que le maximum E se situe à l'âge de 5 ans. On peut donc prendre, en première approximation²¹ $x_0 = 5$ par conséquent $f(x_0) = de = 34,85$. Pour le point d'inflexion C $2x_0 = 10$ et $f(2x_0) = ac = 28,65$.

Il suffit alors de tracer «à main levée», sur du papier quadrillé, une courbe de la forme (4) ayant son maximum au point E de la Figure 3, son inflexion en C , et passant au «plus près» des autres points. En pratique, on peut faire passer la courbe par les points d'abscisse 20, 30, 40 et 50. On trace ensuite la tangente à l'origine – d'où une estimation de $f'(0)$ – et la tangente en C , d'où une estimation de la longueur ab . Nous avons ainsi obtenu (sans tricher) les valeurs approximatives suivantes

$$f'(0) = 21 \quad \text{et} \quad ab = 16$$

d'où

$$\alpha = 16 \quad \beta = 32,308$$

En définitive, les estimations de λ , μ et B , sont en première approximation

²¹ L'unité est l'année en abscisse et représente 100 vivants en ordonnées, si abscisses et ordonnées sont à la même échelle.

$$\hat{\lambda} = \frac{1}{13,63}$$

$$\hat{\mu} = \frac{1}{2,37}$$

$$\hat{B} = 6026$$

On peut comparer ces valeurs à celles données par Lambert.

Que Lambert ait utilisé cette méthode pour obtenir un point de départ à «Des essais nombreux et variés faits à tâtons», comme l'écrivent les P.P. Gaeta et Fontana, on ne le sait pas, et on ne le saura probablement jamais. Mais cela semble hautement vraisemblable et constitue une *deuxième hypothèse* sur la façon dont Lambert a procédé.

La *troisième hypothèse* est la même que celle des P. P. Gaeta et Fontana à partir de sa première approximation, Lambert aurait procédé par tâtonnements – mais pas n'importe lesquels- en jouant sur les paramètres B , λ et μ .

Le Tableau 5 ci-dessous compare aux valeurs observées les résultats de la formule \hat{u}_1 déterminée par la méthode graphique décrite précédemment d'une part, et ceux de la formule \hat{u}_L définitive de Lambert d'autre part. Outre le fait qu'aucune des deux formules ne convient pour l'âge de deux ans, on remarque que \hat{u}_1 sous-estime u pour les âges peu élevés (5, 10 et 20 ans).

On approchera donc mieux les données observées en augmentant B et en diminuant λ : B passe en effet de 6026 à 6176 et λ de $1/13,63$ à $1/13,68$. Le sens des tâtonnements sur μ est moins évident.

Tableau 5. Estimations au moyen de la table X de Süßmilch

x	$\Delta \%$	$\Delta = \hat{u}_1 - u$	1 ^e approx. \hat{u}_1	Observations u	Formule de Lambert		$\Delta \%$
					\hat{u}_L	$\Delta = \hat{u}_L - u$	
2	19	- 606	2 612	3 218	2 623	- 595	18,5
5	1,15	- 40,5	3 445	3 485,5	3 495	+ 9,5	0,3
10	2,1	- 60	2 805	2 865	2 872,5	+ 7,5	0,3
20	2,1	- 29,5	1 388	1 417,5	1 430	+ 12,5	0,9
30	3,3	+ 21,5	668	646,5	689,5	+ 43	6,6
40	13,2	+ 37,5	320,5	283	332	+ 49	17,3
50	5,5	+ 8	154	146	160	+ 14	9,6
60		+ 17,5	73,8	56,3	76,9	+ 20,6	
70		+ 22	35,4	13,5	37	+ 23,5	
80		+ 9,2	17	7,8	17,8	+ 10	
90		+ 8,3	8,2	- 0,1	8,6	+ 8,7	

Ce tableau appelle deux remarques supplémentaires : l'approximation \hat{u}_1 est légèrement meilleure que celle de Lambert pour 30, 40 et 50 ans – ni l'une, ni l'autre des deux formules \hat{u}_1 et \hat{u}_L n'ont de sens au-delà de 50 ans.

Une dernière question se pose – que se serait-il passé si, pour établir sa seconde formule par la méthode hypothétique indiquée ci-dessus, Lambert avait utilisé, non pas les données de la table X de Süßmilch, mais, comme pour sa première formule, celles

que fournissent les registres de mortalité de Londres pour les 6 années 1753 à 1758 incluse ?

En première approximation, on trouve

$$\hat{\lambda} = \frac{1}{13,588} \quad \hat{\mu} = \frac{1}{2,39} \quad \hat{B} = 6158$$

ce qui permet d'établir le Tableau 6. La première approximation \hat{u}_2 est alors meilleure que celle \hat{u}_1 du Tableau 5 pour les premières valeurs de x . Mais la formule de Lambert donne une moins bonne approximation des données – cela n'a rien d'étonnant, puisque les paramètres y ont été estimés à partir d'autres données. Enfin, aucune des deux formules ne s'applique – et ceci à beaucoup près – aux âges supérieurs à 40 ans, de même, bien sûr, qu'à 2 ans et à l'exception – fortuite – de 80 ans.

Tableau 6. Estimations à partir des décès figurant dans les bulletins de mortalité de Londres de 1753 à 1758 (6 années)

x	$\Delta = \hat{u}_2 - u$	Estimation \hat{u}_2	Observations u	Estimations par Lambert \hat{u}_L	$D = \hat{u}_L - u$
2	- 525	2 649	3 174	2 623	- 551
5	+ 2,5	3 503,5	3 501	3 495,5	- 5,5
10	- 1	2 858	2 859	2 872,5	+ 13,5
20	+ 28,5	1 412,5	1 384	1 430	+ 46
30	+ 74	676	602	689,5	+ 87,5
40	+ 121,5	324,5	203	332	+ 129
50	+ 101	155	49	160	+ 111
60	+ 83	74,5	- 8,3	77	+ 85
70		35,7	- 10	37	
80		17,1	18,5	17,8	
90		8,2	4,3	8,6	

Dans le troisième volume de ses *Contributions* Lambert fait d'autres apports importants à l'analyse de la mortalité, qu'il s'agisse du calcul de la «Somme des vivants» ou des conditions d'égalité de la vie moyenne et de la vie probable.

LE CALCUL DE LA «SOMME DES VIVANTS»

Halley et Süßmilch calculaient la somme des vivants en cumulant les décès à partir de la naissance. Ce calcul se justifiait par l'hypothèse, clairement explicitée par l'astronome anglais, que la population était stationnaire. En effet, dans ce cas, la somme des vivants jusqu'à un certain âge correspond à la population totale ayant moins de cet âge.

Lambert calcule, lui, la somme des vivants à partir du dernier âge de la table. Dans celle empruntée selon lui à Süßmilch (cf. Tableau 7²², *infra*), la colonne consacrée à la «Somme des vivants» donne les effectifs d'un âge supérieur à un âge donné :

Les nombres de la quatrième colonne sont de même les sommes des nombres de la troisième : ils représentent la somme des vivants. En effet, si l'on suppose qu'il y a par an 10 000 naissances et 10 000 décès, les nombres de la troisième colonne indiquent combien de personnes vivent en même temps à chaque âge, et les nombres de la quatrième colonne indiquent combien vivent en même temps à chaque âge supérieur²³.

Cette somme des vivants peut également être interprétée, sous la même hypothèse de stationnarité d'une population, comme le nombre de personnes d'âge supérieur à l'âge x . Cette totalisation du nombre des vivants ayant plus d'un certain âge permet aussi un calcul plus direct de la vie moyenne : il suffit de diviser la somme des vivants à l'âge x par le nombre de survivants à cet âge pour obtenir la vie moyenne à l'âge x : ainsi, la vie moyenne à 10 ans est obtenue, dans le tableau 6, en divisant 229 195 par 5538, soit 41,4 années.

Géomètre, Lambert donne une interprétation graphique de la vie moyenne, de la vie probable et des conditions de leur égalité. Il démontre aussi analytiquement ce qu'il a auparavant montré en faisant appel à la géométrie.

L'EGALITE ENTRE VIE MOYENNE ET VIE PROBABLE

En ajoutant ce que Lambert appelle la vie moyenne²⁴ à l'âge x (mais il s'agit en fait de la durée moyenne de survie au-delà de l'âge x) cet âge x , on obtient l'âge moyen au décès²⁵ de ceux qui étaient encore en vie à l'âge x . Par exemple, selon le tableau 7 *supra*, les 5538 personnes encore en vie à 10 ans, ont vécu en moyenne 41,4 + 10 ans, soit 51,4 ans (qui est par conséquent aussi l'âge moyen au décès de ceux encore en vie à 10 ans).

Lambert distingue ce dernier âge de l'âge probable au décès²⁶ (là encore, il s'agit en fait de la durée médiane de survie au-delà de l'âge x). À 10 ans, l'âge probable au décès est la durée de vie qui sera dépassée par la moitié de ceux qui étaient encore en vie à 10 ans. Si 5538 personnes sont encore en vie à l'âge de 10 ans, au bout de combien de temps la moitié, soit 2769, aura-t-elle disparu ? Une interpolation linéaire à partir des chiffres du tableau indique que c'est à 52,4 ans qu'il reste 2769 survivants. C'est l'âge probable au décès de ceux encore en vie à dix ans et leur survie probable à cet âge est donc de 42,4 ans. À cet âge de 10 ans, vie moyenne et vie probable diffèrent donc d'une année. C'est à la naissance et en fin de vie que ces deux résumés de la mortalité diffèrent le plus²⁷.

²² Lambert écrit que cette table (colonnes 1, 2 et 3) provient de Süßmilch. Nous ne l'avons trouvée dans aucune des œuvres publiées de Süßmilch : voici un problème pour les chercheurs !

²³ *Beyträge* III, § 26.

²⁴ «Hittlere Dauer des Lebens».

²⁵ «Hittleres Alter».

²⁶ «Wahrscheinliche Alter». Lambert le nomme aussi : «Alter, wo die Hälfte gestorben [ist]».

²⁷ C'est vrai pour les données du Tableau 7 : écarts maximums à 5 ans et à 81 ans 1/2, écart nul à 30 ans. Mais ce n'est pas une propriété générale des tables de mortalité, ni *a fortiori* une propriété mathématique.

Tableau 7. Table de mortalité de Lambert, tirée d'une table de Süßmilch

Zum §. 24. der Anmerkungen über die Sterblichkeit.

Alter Jahre.	Jähr- lich ster- bende.	Lebende.	Summe der Leben- den.	Es stirbt einer von	Mittlere Alter.	Alter, wo die Halbte ge- storben.
0	2610	10000	295022	4	29,5	22,6
1	610	7390	285022	12	40,2	40,5
2	340	6780	277632	20	42,1	44,4
3	233	6440	270852	27	45,0	46,7
4	169	6207	264412	37	46,7	48,3
5	140	6038	258205	43	47,7	49,5
6	116	5898	252167	51	48,5	50,2
7	96	5782	246269	60	49,6	51,0
8	80	5686	240487	71	50,3	51,6
9	68	5606	234801	82	50,9	52,0
10	58	5538	229195	95	51,4	52,4
11	50	5480	223657	110	51,8	52,8
12	45	5430	218177	121	52,2	53,1
13	42	5385	212747	128	52,5	53,3
14	39	5343	207362	137	52,8	53,6
15	37	5304	202019	143	53,1	53,8
16	35	5267	196715	151	53,3	54,0
17	33	5232	191448	159	53,6	54,2
18	36	5199	186216	144	53,8	54,4
19	39	5163	181017	132	54,1	54,6
20	43	5124	175854	119	54,3	54,8
21	48	5081	170730	106	54,6	55,1
22	54	5033	165649	93	54,5	55,3
23	58	4979	160616	86	55,2	55,6
24	62	4921	155637	79	55,6	55,9
25	65	4859	150716	75	56,0	56,3
26	67	4794	145857	72	56,4	56,7
27	69	4727	141063	69	56,8	57,0
28	70	4658	136336	67	57,3	57,4
29	71	4588	131678	65	57,7	57,8
30	72	4517	127090	63	58,1	58,1
31	74	4445	122573	60	58,6	58,5
32	75	4371	118128	58	59,0	58,9
33	77	4296	113757	56	59,5	59,3
34	79	4219	109461	54	59,9	59,7
35	80	4140	105242	52	60,4	60,1
36	80	4060	101102	51	60,9	60,5
37	81	3980	97042	49	61,4	60,9
38	81	3899	93062	48	61,9	61,3
39	80	3818	89163	48	62,4	61,7
40	80	3738	85345	47	62,8	62,1
41	79	3658	81607	46	63,3	62,5
42	78	3579	77949	45	63,8	62,9
43	77	3501	74370	45	64,2	63,2
44	76	3424	70869	45	64,7	63,6
45	75	3348	67445	44	65,1	64,0
46	75	3273	64097	44	65,5	64,3
47	76	3198	60924	42	66,0	64,7
48	77	3122	57626	41	66,4	65,0
49	78	3045	54504	39	66,9	65,4
50	80	2967	51459	37	67,3	65,8
51	82	2887	48492	35	67,8	66,2
52	84	2805	45605	33	68,2	66,6
53	86	2721	42800	32	68,7	67,0
54	88	2635	40079	30	69,2	67,5
55	90	2547	37444	28	69,7	67,9

55	90	2547	37444	28	69,7	67,9
56	92	2457	34897	27	70,2	68,4
57	94	2365	32440	25	70,7	68,9
58	96	2271	30075	24	71,2	69,4
59	98	2175	27804	22	71,8	70,0
60	100	2077	25629	21	72,3	70,6
61	101	1977	23552	20	72,9	71,2
62	103	1876	21575	18	73,5	71,9
63	105	1773	19699	17	74,1	72,4
64	104	1668	17926	16	74,7	73,1
65	102	1564	16258	15	75,4	73,7
66	99	1462	14694	15	76,0	74,4
67	95	1363	13232	14	76,7	75,1
68	92	1268	11869	14	77,4	75,7
69	88	1176	10601	13	78,0	76,3
70	85	1088	9425	13	78,7	76,9
71	82	1003	8337	12	79,4	77,5
72	80	921	7334	12	80,0	78,1
73	79	841	6413	11	80,6	78,7
74	77	762	5572	10	81,3	79,3
75	75	685	4810	9	82,0	79,9
76	73	610	4125	8	82,7	80,6
77	70	537	3515	8	83,5	81,3
78	66	467	2978	7	84,3	82,0
79	62	401	2511	6	85,2	82,8
80	56	339	2110	6	86,2	83,6
81	50	283	1771	6	87,2	84,5
82	43	233	1488	5	88,4	85,7
83	35	190	1255	5	89,6	87,4
84	25	155	1065	6	90,8	89,3
85	18	130	910	7	92,0	90,9
86	13	112	780	9	92,9	92,0
87	10	99	668	10	93,7	92,8
88	9	89	569	10	94,4	93,0
89	9	80	480	9	95,0	94,1
90	8	72	400	9	95,5	94,7
91	8	64	328	8	96,1	95,3
92	8	56	264	7	96,7	95,9
93	7	48	208	7	97,3	96,5
94	7	41	160	6	97,9	97,1
95	7	34	119	5	98,5	97,7
96	6	27	85	4	99,2	98,3
97	6	21	58	3	99,8	98,9
98	5	15	37	3	100,5	99,5
99	4	10	22	2	101,2	100,3
100	3	6	12	2	102,0	101,0
101	1	3	6	2	103,0	102,5
102	1	2	3	2	103,5	103,0
103		1	1		104,0	

Cette distinction entre vie moyenne et vie probable avait déjà été explicitée par les frères Huygens dans leur correspondance de 1669²⁸, correspondance privée longtemps inconnue, à laquelle Lambert ne pouvait avoir eu accès. Le mathématicien s'appuie par contre sur la présentation que Süßmilch fait de cette distinction dans *L'Ordre divin...* de 1761-62²⁹ à partir des travaux de Halley et de Deparcieux. Lambert utilise la même expression de durée moyenne de la vie (*Die mittlere Dauer des Lebens*) pour désigner ces deux indicateurs de mortalité mais il différencie clairement «la méthode de Halley» (référence à la vie probable) de «la méthode de Deparcieux» (référence à la vie moyenne).

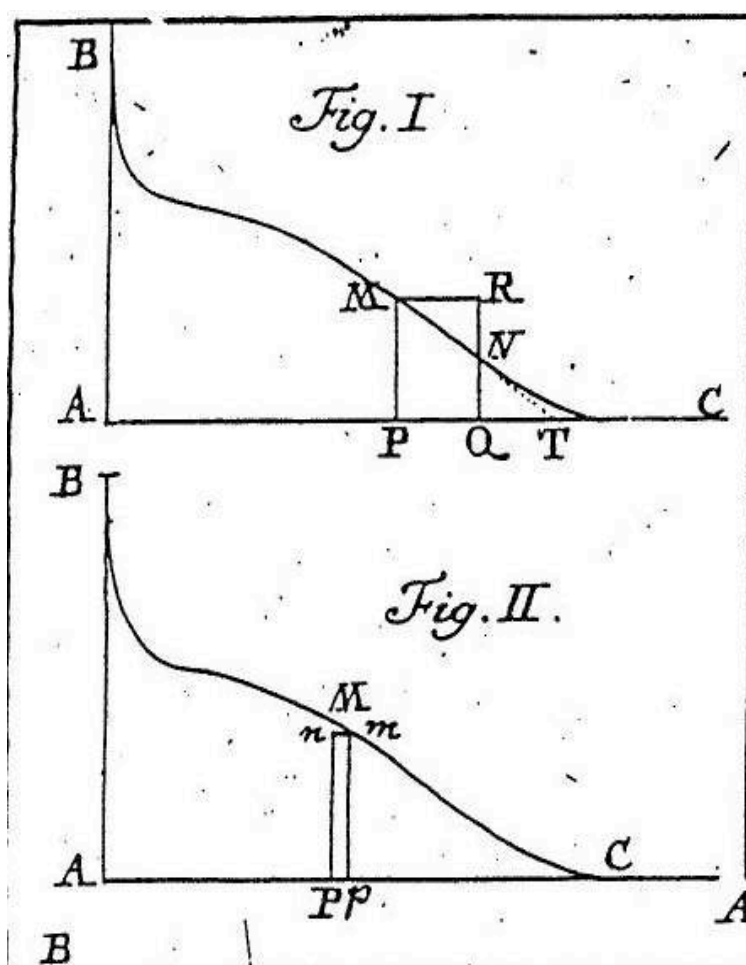


Figure 6. Représentation géométrique par Lambert du calcul des vies moyenne et probable

²⁸ Jacques Véron et Jean-Marc Rohrbasser, «Jodewijk et Christiaan Huygens» la distinction entre vie moyenne et vie probable», *Mathématiques et sciences humaines*, n° 149, 2000, p. 7-21.

²⁹ Süßmilch, *Die göttliche Ordnung...*, 1762, II. XXII, § 473-476, p. 339-348.

Dans ses *Contributions à l'utilisation des mathématiques et à leur application*, Lambert donne une représentation géométrique du mode de calcul des vies moyenne et probable. («Figure II» dans Figure 6). BMC étant la courbe de survie, le point d'abscisse Q représente l'âge auquel survivent la moitié des personnes encore en vie à l'âge P. La vie probable est, par définition, telle que

$$QN = \frac{1}{2} PM$$

Quant à la vie moyenne de ceux qui atteignent l'âge P, elle est donnée par

$$\frac{PCM}{PM}$$

et l'âge moyen au décès de ceux-ci correspond alors :

$$AP + \frac{PCM}{PM}$$

La condition d'égalité entre âge probable et âge moyen au décès est alors

$$AQ = AP + \frac{PCM}{PM}$$

ou encore

$$PQ = \frac{PCM}{PM}$$

Il faut par conséquent que

$$PQ \times PM = PCM$$

Étant donné que $PM \times PQ$ donne la surface du rectangle PMRQ, la contrainte d'égalité des âges moyen et probable implique que la surface PCM soit égale à la surface PMRQ. Il y a cette égalité lorsque la courbe de survie est une droite³⁰, comme le démontre Lambert à partir d'un ajustement numérique de la forme³¹

$$x = ay^n + by^m + \text{etc.}$$

dont, pour des exposants entiers, la seule solution est

$$x = a + by$$

Ce résultat implique que, dans l'hypothèse où la courbe $x(y)$ serait une parabole, l'égalité entre âge moyen et âge probable est impossible. Par contre, Lambert montre que, dans le cas où la fonction de survie est du type $y = ax^q$, il y a une proportionnalité entre ces deux âges. La proportion dépend de la valeur de q (sur toute question de l'égalité entre vies moyenne et probable, se reporter à l'Annexe V).

³⁰ Ce qui reviendrait à une égalité entre les triangles MNR et NQC.

³¹ Dans cette équation, x est l'écart à l'âge maximum et y est le nombre correspondant de survivants. Lambert considère donc la fonction $x(y)$ inverse de $y(x)$.

LA «FORCE VITALE»

Pour Lambert le risque de décéder à chaque âge permet «la mesure réelle de la *force vitale* ».

Inverse de la probabilité de décéder, elle est exprimée en années, par la sous-tangente en chaque point de la courbe de survie. Ainsi, à l'âge de 1 an, cette force vitale atteindrait 12 ans ; elle serait, à l'âge de 17 ans, maximale avec une valeur de 159 ans (cf. Tableau 7, ci-dessus).

Selon Lambert, non seulement la force vitale ne diminuerait pas aux grands âges, mais elle pourrait même augmenter légèrement :

*Elle croît très rapidement jusqu'à la 16^e ou 17^e année. A partir de là, elle décroît également très rapidement jusqu'à la 25^e année, ensuite plus lentement, et, à partir de la 50^e année de nouveau un peu plus rapidement. Au milieu de la quatre-vingtième année, elle semble au contraire quelque peu croître à nouveau ; ainsi, dans la 95^e année, elle n'est pas plus petite que dans la 83^e.*³²

Le mathématicien s'interroge sur la causalité en jeu dans les variations de la force vitale. Une évolution régulière de celle-ci suggère qu'une cause déterminante est à l'œuvre :

*Si, jusqu'à la 15^e année environ, la force vitale s'accroissait entièrement suivant une progression arithmétique, on pourrait en conclure que cela devrait avoir une cause efficiente tout à fait simple et uniforme. [...] Or, il ne manque pas grand chose pour que la progression arithmétique que nous avons signalée soit régulière, et il semble que au moins la cause principale de l'augmentation de la force vitale doive être simple.*³³

La cause déterminante serait contrariée, entre les âges de 17 et 25 ans, car la force qui devrait à ses yeux continuer de croître, en fait diminue. Tout se passe pour lui, qui aime les comparaisons d'ordre physique, comme dans le cas d'«Un cylindre contenant un siphon qui ne commence à vider le cylindre que lorsque ce dernier est rempli à un niveau plus élevé que celui qu'atteint le siphon». La loi de survie est pour Lambert du même ordre que tout phénomène naturel.

Cette discussion sur les variations de la «force vitale» repose sur l'observation des valeurs fournies par la table empirique de Süßmilch (cf. Annexe I).

À partir de la seconde formule théorique de Lambert :

$$y = 10000 \left(\frac{96-x}{96} \right)^2 - 6176 \left(e^{-\frac{x}{13,682}} - e^{-\frac{x}{2,43114}} \right)$$

on remarquera que la force vitale, $F(x)$ est, par définition, le rapport :

$$F(x) = -\frac{y'(x)}{y''(x)} = \frac{y}{|y'|}$$

³² *Beyträge* III, § 48.

³³ *Beyträge* III, § 51.

Entre les deux points d'inflexion de la courbe $x'_1 \# 10,67$ et $x_1 \# 37,23$ (cf. Annexe V), $|y'|$ est monotone croissante. Comme y est une fonction décroissante, $F(x)$ ne peut que décroître dans un intervalle qui inclut celui qui s'étend de 17 à 25 ans.

Pourquoi Lambert ne l'a-t-il pas observé?

*
* *

Dans ses *Contributions à l'utilisation des mathématiques et à leur application*, Lambert aborde en mathématicien les questions de population — il fait autant appel à l'analyse qu'à la géométrie. Il se montre tout à fait conscient du caractère contingent des résultats numériques obtenus — leur validité dépend étroitement du nombre et de la qualité des données de mortalité dont il dispose.

Pour Lambert le monde est organisé selon des régularités, la nature n'y fait pas de «saut». De fortes discontinuités dans des séries de nombres sont incompatibles avec l'ordre de la nature, ce qui conduit parfois le mathématicien à lisser des courbes manifestement irrégulières. Sa démarche est par ailleurs celle d'un expérimentateur — il propose des «exemples» qui ont valeur de modèles.

Une des vertus des *Contributions à l'utilisation des mathématiques et à leur application* tient à la rigueur de la démarche adoptée par Lambert alliée à une grande attention portée au réel. On retiendra tout particulièrement de cet ouvrage une formalisation originale et pionnière, un demi-siècle avant Gompertz, de la loi de mortalité. Avec son contemporain Euler, Lambert peut être considéré comme un des fondateurs de la démographie mathématique.

ANNEXE II

Sur la première formule de Lambert

Si x désigne l'âge moins 45 ans, le nombre des survivants est, selon la première formule de Lambert

$y(x)$ = nombre de survivants au temps x

$$= a - bx + cx^2 - dx^3 - ex^4 + fx^5$$

avec

$$a = 26.950$$

$$d = 0,03427$$

$$b = 987,5$$

$$e = 0,0027017$$

$$c = 9,7091$$

$$f = 0,000066635$$

Les dérivées successives de $y(x)$ sont

$$y'(x) = -b + 2cx - 3dx^2 - 4ex^3 + 5fx^4$$

$$y''(x) = 2(c - 3dx - 6ex^2 + 10fx^3)$$

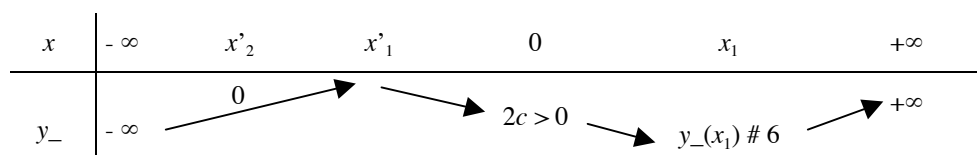
$$y'''(x) = 6(-d - 4ex + 10fx^2)$$

y''' , (trinôme du second degré), a deux racines x'_1 et x_1 que l'on sait calculer, et qui valent approximativement

$$x'_1 \# -2,716$$

$$x_1 \# 18,934$$

D'où, pour les variations de la cubique y'' , qui vérifie $y''(0) = 2c > 0$



Le minimum $y''(x_1)$ vaut approximativement 6, valeur positive.

Il en résulte que pour $x \geq 0$, le polynôme du 5^e degré $y(x)$ a sa concavité toujours dans le même sens, et que sa dérivée $y'(x)$ est monotone croissante.

On vérifie que

$$y'(44) \# -3,934 < 0$$

$$y'(44,5) \# 27,22 > 0$$

En améliorant cet encadrement, on obtient, pour le zéro x_0 de y' , et par suite pour le minimum de y

$$x_0 \# 44,06$$

d'où

$$y(x_0) \# 240,403$$

L'allure de la courbe représentative de $y(x)$ est donnée par la Figure II.1.

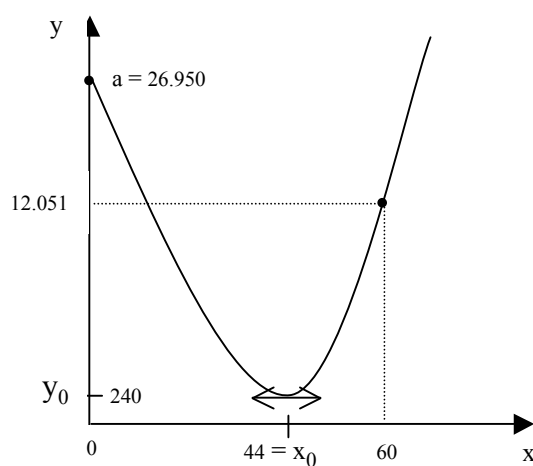


Figure II.1

Il en résulte que la formule n'a de sens que jusqu'à $x = 44$, i.e. un âge de 89 ans.

La valeur moyenne m de x se calcule aisément au moyen de l'intégrale $\int_0^{44} x dy$, soit : $m \approx 17$ (62 ans).

Quant à l'âge médian μ , il vaut approximativement :

$$\mu = 45 + 16 = 61 \text{ ans.}$$

Bien entendu, ces valeurs s'entendent pour les survivants à 45 ans.

ANNEXE III

Seconde formule de Lambert : inflexions, vie moyenne

Pour 10.000 naissances, le nombre $y(x)$ de survivants à l'âge x est fourni par l'expression :

$$(0) \quad y(x) = A \left(\frac{(a-x)}{a} \right)^2 - B(e^{-\lambda x} - e^{-\mu x})$$

$$\text{avec : } A = 10.000, \quad B = 6.176, \quad a = 96, \quad \lambda = \frac{1}{13,682}, \quad \mu = \frac{1}{2,43114}$$

1. Variations de y et points d'inflexion.

Les dérivées successives de $y(x)$ sont :

$$(1) \quad y'(x) = \frac{-2A(a-x)}{a^2} - B(-\lambda e^{-\lambda x} + \mu e^{-\mu x})$$

$$(2) \quad y''(x) = \frac{2A}{a^2} - B(\lambda^2 e^{-\lambda x} - \mu^2 e^{-\mu x})$$

$$(3) \quad y'''(x) = -B(\mu^3 e^{-\mu x} - \lambda^3 e^{-\lambda x})$$

De (3) on déduit que y''' s'annule pour (l_n désigne le log-neperien) :

$$x_0 = 3 \frac{l_n(\mu/\lambda)}{\mu - \lambda} \# 15,324$$

$$\text{Comme : } y''(0) = \frac{2A}{a^2} + B(\mu^2 - \lambda^2) > 0,$$

$$y''(x_0) \# -23 < 0 \quad \text{et} \quad y''(\infty) = \frac{2A}{a^2} > 0$$

$y''(x)$ s'annule pour deux valeurs x'_1 et x_1 (les abscisses des points d'inflexion).

D'après (2), en posant $e^{-\lambda x} = Z$ et $\frac{\mu}{\lambda} = v$, et en itérant la transformation :

$$Z \mapsto v^2 Z^v + \frac{2A}{a^2 B \lambda^2}$$

on obtient des approximations successives de $Z'_1 = e^{-\lambda x'_1}$ qui en est un point fixe attractif.

On procède de même pour x_1 , au moyen de la transformation :

$$Z \mapsto \frac{1}{v^2} Z^{1/v} - \frac{2A}{a^2 B \mu^2}$$

On obtient ainsi les approximations :

$$x'_1 \# 10,6768 \quad x_1 \# 37,231$$

d'où, d'après (1), les pentes aux points d'inflexion :

$$y'(x'_1) \# -14 \text{ (i.e., quasi-horizontalité)}$$

$$y'(x_1) \# -97$$

On peut alors dresser le tableau des variations des dérivées de $y(x)$ et par suite de $y(x)$ lui-même :

x	0	x'_1	x_0	x_1	x_3	x_2	x'_3	$+\infty$			
y'''	—	—	—	0	+	+	+	+			
y''		0		-23		0					
y'		-14		-33		-97		0		$+\infty$	
y	$A>0$	5.146		3.341		0		-5,27		0	$+\infty$

y' s'annule pour une valeur $x_2 \# 95,811$ (on a $y'(95) \# -1,73$ et $y'(96) \# 0,04$).

Le minimum correspondant de y est :

$$y(x_2) = -5,27$$

C'est dire (il faut le rapporter à 10.000) qu'il est très proche de zéro.

Néanmoins, $y(x)$ va s'annuler pour deux valeurs x_3 et x'_3 , avec :

$$x_3 \# 93,525 \quad \text{et} \quad 98 < x'_3 < 99.$$

Finalement, c'est x_3 , c'est-à-dire à peu près 93 ans et 6 mois, et non 96 ans, qui est l'âge maximum compatible avec la formule de Lambert.

Compte tenu de l'ordre de grandeur du minimum négatif de $y(x)$, considérer 96 ans comme l'âge maximum ne change pas grand-chose.

Retenons en définitive que les deux inflexions de la courbe de mortalité de Lambert se situent approximativement à 10 ans et 8 mois pour le premier, et 37 ans et 3 mois pour le second.

2. Calcul de la vie moyenne à l'âge x

Elle s'obtient au moyen de l'intégrale (le signe « - » résulte de ce que $y(x)$ et décroissante)

$$(4) \quad I(X) = \int_0^X x dy = B \left(\frac{1}{\lambda} - \frac{1}{\mu} - \left(X + \frac{1}{\lambda} \right) e^{-\lambda x} + \left(X + \frac{1}{\mu} \right) e^{-\mu x} \right) + \frac{2A}{a^2} X^2 \left(\frac{X}{3} - \frac{a}{2} \right)$$

L'âge maximum impliqué par la formule (0) de Lambert étant x_3 , voisin de 93 ans 6 mois, l'espérance de vie $m(x)$ à l'âge x est :

$$(5) \quad m(x) = - \frac{1}{y(x)} \int_x^{x_3} t dy = \frac{I(x) - I(x_3)}{y(x)}$$

Notamment, l'espérance de vie à la naissance est, selon la formule de Lambert, de 25 ans environ.

Le calcul des durées moyennes de survie $s(x) = m(x) - x$ à quelques autres âges fournit le Tableau III.1 suivant :

x	$s(x)$
1	29,55
2	33,65
3	35,5
4	37,95
5	39,12

x	$s(x)$
10	36,8
15	32,4
20	28,8
30	23,45
40	19,23

x	$s(x)$
50	15,15
60	12
70	8,6
80	5
90	2,95

Tableau III.1

On pourra comparer ces valeurs *théoriques* aux valeurs empiriques fournies par les tables de mortalité contemporaines (Süßmilch, Deparcieux, etc...) et notamment à celles du Tableau 7 ci-dessus.

ANNEXE IV

Etude des fonctions $f(x) = B(e^{-\lambda x} - e^{-\mu x})$

On suppose, comme c'est le cas pour la formule de Lambert :

$$\mu > \lambda > 0, \quad B > 0$$

Dans la suite, on posera :

$$\frac{\mu}{\lambda} = \nu > 1 \quad \text{et} \quad \nu^{\frac{1}{\nu-1}} = \Pi$$

On a :

$$(1) \quad f(x) = B(e^{-\lambda x} - e^{-\mu x})$$

D'où pour les dérivées :

$$(2) \quad f'(x) = B(\mu e^{-\mu x} - \lambda e^{-\lambda x}) = B\lambda e^{-\mu x}(\nu - e^{(\mu-\lambda)x})$$

$$(3) \quad f''(x) = B(\lambda^2 e^{-\lambda x} - \mu^2 e^{-\mu x}) = B\lambda^2 e^{-\mu x}(e^{(\mu-\lambda)x} - \nu^2)$$

De (2) on déduit :

$$(4) \quad f'(0) = B(\mu - \lambda)$$

et que $f'(x)$ s'annule pour :

$$x_0 = \frac{\ln \nu}{\mu - \lambda}$$

D'où : $-\lambda x_0 = \ln \Pi$ et $-\mu x_0 = \ln(\Pi)^\nu$

On en déduit, pour le maximum de la fonction f :

$$(5) \quad f(x_0) = B(\Pi - \Pi^\nu) = B \Pi \left(1 - \frac{1}{\nu}\right)$$

De même, f'' s'annule pour :

$$x_1 = \frac{\ln \nu^2}{\mu - \lambda} = 2 \frac{\ln \nu}{\mu - \lambda} = 2x_0$$

Et l'on obtient, pour l'ordonnée et la pente au point d'inflexion :

$$(6) \quad f(2x_0) = B(\Pi^2 - \Pi^{2\nu}) = B \Pi^2 \left(1 - \frac{1}{\nu^2}\right)$$

$$(7) \quad f'(2x_0) = B(\mu \Pi^{2\nu} - \lambda \Pi^2) = -B \Pi^2 \left(1 - \frac{1}{\nu}\right) < 0$$

Le tableau des variations de f est ainsi :

x	0	x_0	$2x_0$	$+\infty$
f''	—	—	0	+
f'	$3(\mu - \lambda) > 0$	$\rightarrow 0$	$\rightarrow f'(2x_0)$	$\rightarrow 0$
f	0	$\rightarrow f(x_0)$	$\rightarrow f(2x_0)$	$\rightarrow 0$

D'autre part, on déduit de (6) et (7) :

$$(8) \quad \frac{f(2x_0)}{f'(2x_0)} = -\frac{1}{\lambda} \left(1 + \frac{1}{\nu}\right) = -\frac{1}{\lambda} \left(1 + \frac{\lambda}{\mu}\right) = -\left(\frac{1}{\lambda} + \frac{1}{\mu}\right)$$

Puis de (4), (5) et (6) :

$$(9) \quad \frac{f'(0) \cdot f(2x_0)}{f^2(x_0)} = \frac{B(\mu - \lambda) \cdot B \Pi^2 \left(1 - \frac{1}{\nu^2}\right)}{B^2 \Pi^2 \left(1 - \frac{1}{\nu}\right)^2} = (\mu - \lambda) \frac{1 + \frac{1}{\nu}}{1 - \frac{1}{\nu}}$$

$$= (\mu - \lambda) \frac{\mu + \lambda}{\mu - \lambda} = \mu + \lambda$$

On a vu plus haut comment, de (8) et (9) on peut déduire des estimations de λ et μ .

ANNEXE V

Egalité et proportionnalité entre vie moyenne et vie probable

1. Comme le fit Lambert pour traiter ces questions (§ 36 à 45, p. 501-508 de son traité), nous utiliserons ici
 - 1.1. la fonction $y(x)$, où x est l'écart de l'âge considéré à la durée c maximum de la vie, et $y(x)$ le nombre correspondant de vivants.
 - 1.2. $y(x)$ est donc monotone non décroissante, de 0 à y_0 (y_0 est le nombre de vivants à la naissance) lorsque x varie de 0 à c : $y(0) = 0$, $y(c) = y_0$. La courbe représentative de $y(x)$ est d'ailleurs la symétrique de la courbe de mortalité habituelle, par symétrie axiale par rapport à la droite d'équation $x = \frac{c}{2}$ (cf. Figure V.1).

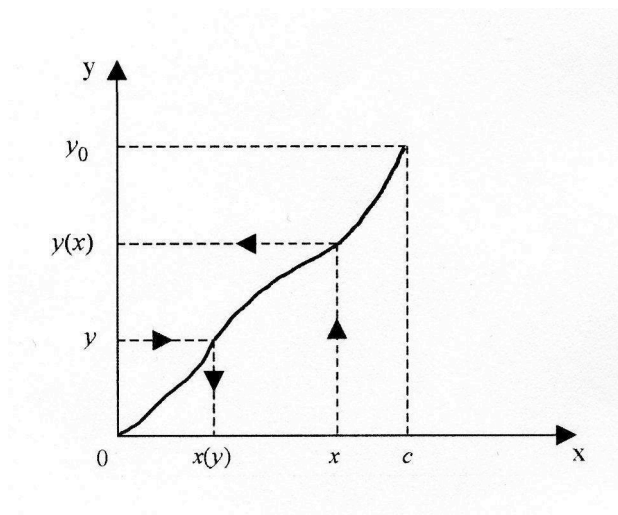


Figure V.1

- 1.3. Nous supposons en outre (c'est implicite chez Lambert) $y(x)$ monotone strictement croissante et continue, de façon à assurer l'existence de la fonction $x(y)$ inverse de $y(x)$.
- 1.4. Enfin, comme, dans la suite, des raisonnements et des calculs se font, comme chez Lambert, sur cette fonction inverse $x(y)$, nous introduirons, en tant que de besoin, les hypothèses supplémentaires qui sont implicites dans les calculs de Lambert, mais qu'il vaut mieux expliciter.
En particulier, $x(y)$ qui est nécessairement continue et monotone strictement croissante, est en outre supposée dérivable dans l'intervalle $[0, y_0]$, avec $x'(y) > 0$.
- 1.5. Là où la démarche qui suit diffère un peu de celle de Lambert, c'est que nous allons traiter simultanément la question de la proportionnalité entre vie moyenne et vie probable (i.e., médiane) et celle de l'égalité, cette dernière étant un cas particulier de la première.

2. L'équation différentielle caractéristique.

Soit $m(x)$ la durée moyenne de survie à l'âge $(c - x)$. Par définition, et compte tenu des conventions 1.1 et 1.2 ci-dessus, on a :

$$(1) \quad m(x) = \frac{1}{y(x)} \int_0^x t dy(t), \text{ soit encore } \square$$

$$(1') \quad m(x)y(x) = \int_0^x t dy(t)$$

De même, pour la survie probable (médiane) $\mu(x)$ \square

$$(2) \quad \forall x, y[\mu(x)] = \frac{1}{2} y(x)$$

soit, de façon équivalente \square

$$\mu = x \left(\frac{y}{2} \right) \quad (\forall y \in [0, y_0])$$

Enfin, si l'on suppose μ proportionnel à m , et en désignant par $\lambda(> 0)$ le coefficient de proportionnalité \square

$$(3) \quad \forall x, \mu(x) = \lambda m(x)$$

Portons (3) dans (1') \square il vient \square

$$(4) \quad \mu(x) y(x) = \lambda \int_0^x t dy(t)$$

Différentions les deux membres de (4) \square

$$(4') \quad \mu dy + y d\mu = \lambda x dy$$

Or, d'après (2') \square

$$\mu = x \left(\frac{y}{2} \right) \text{ et donc } d\mu = x' \left(\frac{y}{2} \right) \cdot \frac{dy}{2}$$

Il vient alors \square

$$\lambda x(y) = x \left(\frac{y}{2} \right) + \frac{y}{2} x' \left(\frac{y}{2} \right)$$

Soit en posant $\square \frac{y}{2} = z \quad (0 \leq z \leq \frac{y_0}{2}) \square$

$$(5) \quad \lambda x(2z) = x(z) + zx'(z) = \frac{d}{dz}(zx(z))$$

Toute solution dérivable du problème posé doit être solution de l'équation différentielle linéaire (5).

D'autre part, si l'on intègre «par parties» le second membre de (4), on obtient

$$(4'') \quad \frac{1}{\lambda} y(x) \mu(x) = xy(x) - \int_0^x y(t) dt$$

$y(x) \mu(x)$, c'est la surface du rectangle $0\mu Hy$ (cf. Figure V.2)

$xy(x)$, c'est la surface du rectangle $0xMy$

$\int_0^x y(t) dt$, c'est la surface du «triangle» curviligne $0xM$.

Le second membre de (4'') est donc la surface du «triangle» curviligne $0My$.

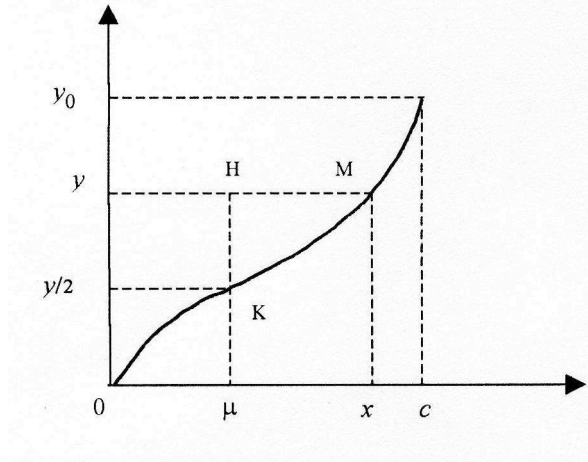


Figure V.2

Dans le cas où $\lambda = 1$, on doit donc avoir égalité entre les surfaces des deux «triangles» curvilignes $0\mu K$ et KMH .

C'est ce que Lambert avait bien remarqué, comme c'est rappelé plus haut (cf. p. 22, et sa Figure I, p. 21). Evidemment, une condition suffisante pour cette égalité (quel que soit x), c'est que la courbe de mortalité soit un segment de droite (cf. Figure V.3)

3. Solutions de (5) parmi les fonctions $x(z)$ entières

Revenons à l'équation différentielle (5) et cherchons ses solutions de la forme

$$(6) \quad x(z) = a_0 + a_1 z + a_2 z^2 + \dots = \sum_{n=0}^{\infty} a_n z^n$$

la série du second membre étant absolument et uniformément convergente dans l'intervalle $0 \leq z \leq y_0$. Selon son texte (§ 39, p. 502-503 et p. 67 ci-dessus), il semble que Lambert se soit borné au cas où $x(z)$ est un polynôme.

D'après (5), on doit avoir, quel que soit z , et avec $\lambda > 0$:

$$(5') \quad \lambda(a_0 + 2a_1 z + 4a_2 z^2 + \dots + z^n a_n z^n + \dots = \\ a_0 + 2a_1 z + 3a_2 z^2 + \dots + (n+1)a_n z^n + \dots$$

Les deux membres de (5') doivent être identiques terme à terme.

Or on a la propriété bien connue

$$\forall n \geq 2 \quad 2^n > n+1 \quad 2^0 = 1 \quad 2^1 = 1+1$$

Il en résulte que (5') n'a aucune solution pour $\lambda > 1$.

Pour $\lambda = 1$, on a la solution

$$x(x) = a_0 + a_1 z$$

comme en outre $x(0) = 0$, cette solution est $x(z) = a_1 z$, et finalement

$$y(x) = ax$$

c'est la solution linéaire fournie par Lambert (cf. p. 137) et évoquée ici à la fin du § 2. On peut l'appeler «courbe de mortalité de Leibniz» (cf. Leibniz et les raisonnements sur la vie humaine, INED, Paris, 2001)

Pour $0 < \lambda < 1$ (rappelons que, nécessairement, $\lambda > 0$), certaines valeurs de λ fournissent chacune une solution de (5). Par exemple, $\lambda = \frac{3}{4}$ donne

$$x(z) = Az^2 \quad (\text{avec } A > 0, \text{ car } x(z) \text{ doit être monotone croissante}).$$

De même, $\lambda = \frac{1}{2}$ fournit $x(z) = Az^3$, $A > 0$.

D'une façon générale, si, pour un entier n positif, on a

$$\lambda = \lambda_n = (n+1) 2^{-n}$$

on aura la solution $x(z) = Az^n$ ($A > 0$)

D'où pour $y(x)$ une expression de la forme

$$(6) \quad y = ax^q \quad (a > 0 \quad q = \frac{1}{n})$$

En fonction de $q = \frac{1}{n}$, le coefficient de proportionnalité λ est

$$\lambda = \frac{q+1}{q} 2^{-1/q}$$

4. Solutions de (5) parmi les fonctions puissance

Ceci nous amène à rechercher les solutions de (5) de la forme

$$x(z) = Az^\alpha$$

avec $A > 0$ et $\alpha > 0$ puisque $x(z)$ doit être monotone strictement croissante. α est maintenant un réel quelconque, et non nécessairement entier.

Comme $x'(z) = A\alpha z^{\alpha-1}$

On a, d'après (5)

$$\lambda 2^\alpha = 1 + \alpha, \text{ soit } \lambda = (1 + \alpha)2^{-\alpha}$$

On va donc étudier la fonction f d'une variable réelle t définie par

$$f(t) = (1 + t)2^{-t}$$

et chercher les solutions de l'équation $f(t) = \lambda$, pour $\lambda > 0$.

C'est-à-dire $t > -1$

Le calcul des dérivées première et seconde de f montre que

$$f \text{ s'annule en } t_0 = \frac{1 - \ln 2}{\ln 2} \approx 0,4427$$

d'où pour le maximum de $f(t)$

$$f(t_0) = \lambda_0 = \frac{2}{\ln 2} 2^{-1/\ln 2} \approx 1,0615$$

$$f' \text{ s'annule pour } t_1 = \frac{2}{\ln 2} - 1 \approx 1,8854$$

d'où pour l'inflexion de $f(t)$

$$f(t_1) = \frac{4}{\ln^2 2} \left(2^{-1/\ln 2} \right)^2 \approx 0,781$$

$$f''(t_1) = -2^{-t_1} \approx -0,27$$

On a ainsi le tableau de variation

t	-1	0	t_0	1	t_1	2	$+\infty$
f'	2	+	0	-	-	-2^{-t_1}	-
f	0	1	λ_0	1	$\frac{3}{4}$	0	

Appelons (C) la courbe représentative de f . Elle est continue, et a pour asymptote l'axe horizontal.

On peut maintenant examiner tous les cas possibles.

$$0 < \lambda < 1$$

On obtient deux intersections de la courbe (C) avec la droite horizontale d'ordonnée λ ; soient α et α' leurs abscisses. On a :

$$-1 < \alpha' < 0 \quad \text{et} \quad 1 < \alpha$$

D'où les solutions :

$$x = Az^\alpha, \quad \text{i.e.} \quad y = ax^q, \quad \text{avec} \quad 0 < q = \frac{1}{\alpha} < 1$$

et
$$x = A'z^{\alpha'}, \quad \text{i.e.} \quad y = a'x^{q'}, \quad \text{avec} \quad q' = \frac{1}{\alpha'} < -1$$

La seconde, qui est une courbe de Pareto, ne convient pas, puisque $x(z)$ et $y(x)$ doivent être monotones croissantes.

La première (cf. Figure V.4) correspondrait à une courbe de mortalité dans laquelle il y aurait accélération de la mortalité aux grands âges ; plus généralement, le nombre annuel de morts croîtrait constamment avec l'âge. Est-ce réaliste ?

Les solutions trouvées § 3 supra sont parmi celles-ci.

$$\lambda = 1$$

$f(t) = 1$ a pour solutions $t = 0$ et $t = 1$.

La seconde correspond au cas déjà vu d'une courbe de mortalité linéaire (cf. Figure V.3). C'est la « courbe de mortalité de Leibniz » le nombre de morts est le même à tous les âges.

Quant à la première ($t = 0$), elle fournit la solution singulière

$$x = Az^0 = A$$

Si x est constant, toute la population a le même âge aucune « pyramide des âges » n'a cette propriété.

$$1 < \lambda < \lambda_0$$

La droite horizontale d'ordonnée λ coupe la courbe (C) en deux points d'abscisses α et α' , avec

$$0 < \alpha < t_0 < \alpha' < 1$$

d'où, pour $y(x)$, deux fonctions puissance

$$y = ax^q \quad \text{et} \quad y = a'x^{q'}$$

avec

$$\frac{1}{\alpha} = q > \frac{1}{t_0} = q_0 > \frac{1}{\alpha'} = q' > 1$$

Dans son texte (§ 44, p. 505-507) Lambert n'envisage pour l'exposant q , que les valeurs entières 1, 2, 3, 4, ...

Les solutions qu'il fournit entrent donc dans cette catégorie. Il remarque d'ailleurs bien que si l'exposant q augmente, le rapport λ de la survie probable à la survie moyenne décroît et tend vers un.

$$\lambda = \lambda_0$$

On a la «solution double»

$$y = ax^{q_0} \quad \text{avec} \quad q_0 = \frac{1}{t_0} \approx 2,3422$$

La Figure V.5 donne l'allure, plus réaliste pour le XVIII^e siècle que la précédente, d'une courbe de mortalité correspondant aux cas $1 < \lambda \leq \lambda_0$

Remarque finale. Pour $\lambda = 1$, l'équation différentielle (5) s'écrit encore

$$(5') \quad x'(z) = \frac{x(2z) - x(z)}{z}$$

Ceci signifie qu'en chaque point, d'abscisse z , la tangente à la courbe représentative de $x(z)$ est la corde qui joint ce point à celui d'abscisse $2z$ (cf. Figure V.6).

Cette observation rend intuitif le fait que cette courbe soit un segment de droite.

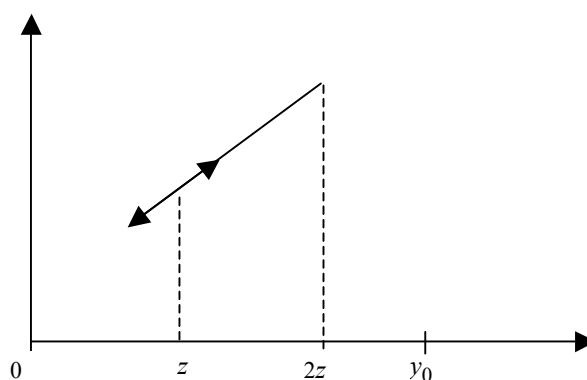


Figure V.6

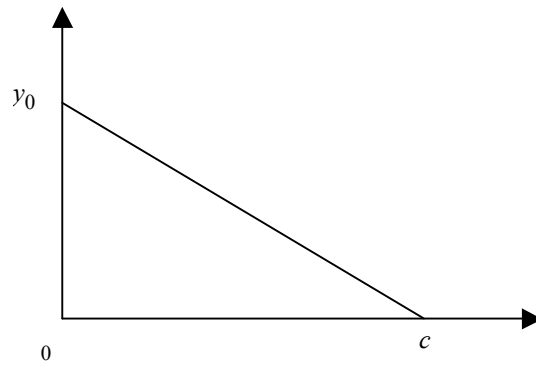


Figure V.3. $\lambda = 1$

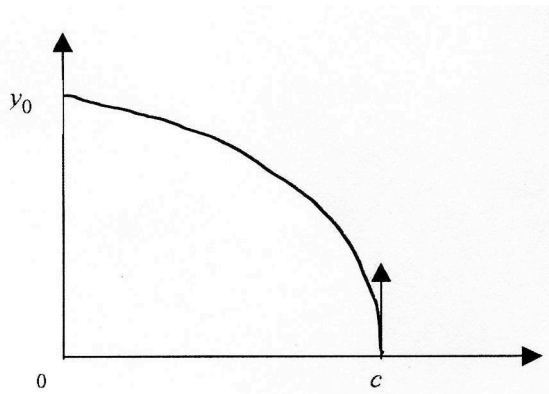


Figure V.4. $0 < \lambda < 1$

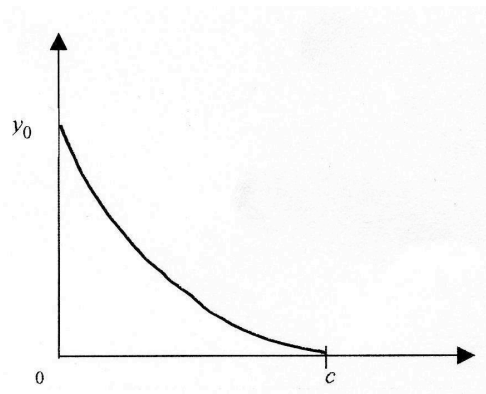


Figure V.5. $1 < \lambda \leq \lambda_0$

On remarquera qu'aucune de ces courbes n'a de point d'inflexion.